DOCUMENT RESUME

ED 238 916 TM 840 006

AUTHOR	Boruch, Robert F.; Wortman, Paul M.
TITLE	Project on Secondary Analysis, 1979-82. Final Report. Report No. A-95.
INSTITUTION	Northwestern Univ., Evanston, Ill. Dept. of Psychology.
SPONS AGENCY	National Inst. of Education (ED), Washington, DC.
PUB DATE	May 83
GRANT	NIE-G-79-0128
NOTE	31p.; For a related document, see TM 830 004.
PUB TYPE	Reports - Evaluative/Feasibility (142)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	*Data Analysis; *Data Collection; Evaluation Methods; Evaluation Utilization; Policy Formation; *Program Evaluation; *Research Methodology; Research Utilization; Surveys
IDENTIFIERS	*Secondary Analysis

ABSTRACT

The Project on Secondary Analysis directed its attention to methodology and policy issues, focusing on reanalysis of existing data (1979-82)--surveys, program evaluations, etc.--as a vehicle for learning how to improve methods of obtaining and analyzing data, designing program evaluations, and constructing policy to enhance the quality and utility of statistical data. This document contains: (1) an executive summary, (2) an introduction, (3) a summary of project products and their uses, (4) the third year report and abstracts (1981-82), (5) the second year report and abstracts (1980-81), (6) a reference list of journal articles and books, and (7) appendices including reprints of illustrative articles. (PN)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED238916

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

TITLE: Final Report
Project on Secondary Analysis
NIE Crant G-79-0128
1979-82

PRINCIPAL INVESTIGATORS: Robert F. Boruch
and
Paul M. Wortman

DATE: May, 1983.

PREPARED FOR. Dan Antonoplos, Project Monitor, NIE

REPORT NO.: A-95

INSTITUTION: Department of Psychology
Northwestern University
Evanston, Illinois 60201

Contents

I. Executive Summary

II. Introduction

III. Products and their Uses

IV. Third Year Report and Abstracts: 1981-82

V. Second Year Report and Abstracts: 1980-81

VI. Reference List of Journal Articles and Books

VII. Appendices: Reprints of Illustrative Articles

I. Executive Summary
Project on Secondary Analysis

1. NIE support during 1979-82 led to research that was eventually published in three books, over 40 articles published in scholarly journals and four dissertations, and about 30 presentations at professional meetings and conferences. Nearly 15 graduate students received full or partial support for research studies.

2. There is substantial evidence that the results of this research have been used in agencies such as GAO, the Department of Education, National Academy of Sciences, Social Science Research Council, and UNESCO. Further, the work is frequently cited in the professional literature. The American Educational Research Association gave its "Research Review Award" to our paper based on the research in 1981.

3. The methodological products include better specification of conditions under which randomized field experiments are feasible and appropriate for evaluating education programs, analytic solutions to problems engendered by the use of regression-discontinuity designs for program evaluation and an array of technical solutions to assuring privacy and confidentiality in research.

4. The published policy products include coherent guidelines and supporting evidence on (a) evaluation policy at the federal and state level (b) policy on enhancing access to and utility of large-scale surveys such as NAEP and High School and Beyond, (c) policy on access to program evaluation data in particular and scientific research data in general.

5. Our published research on use of educational research counters prevailing views that such data are not used, documents use at the local, state and federal levels of some 40 large-scale program evaluations and expensive survey research efforts such as NAEP and High School and Beyond. Recommendations for enhancing the livelihood of data use have been made and considered by policy groups that govern such studies.

6. The published reanalyses of data sets sometimes confirmed earlier findings, as in reanalysis of large-scale Colombian compensatory Education experiments. They disconfirmed some findings in initial analyses of the national Irish experiments on achievement testing. They led to better understanding of the structure of mental abilities and their link to age and special subpopulations, e.g., Indians, Samoans, Hispanics in the case of IQ tests.

7. Our work on solutions to ethical and legal problems in research led to state of the art analyses of costs and benefits of technical statutory and procedural methods for assuring privacy, and for assuring equity in randomized field experiments.

II. Introduction

This is the final report of our NIE supported Project on Secondary Analysis.

It summarizes work done between 1979-82, identifies the journals that report the research, and outlines some uses to which the data have been put.

The director of the Project was Robert F. Boruch, Northwestern University. Paul Wortman, now at University of Michigan was co-director.

Others involved in research included William Trochim, now at Cornell, Joe Locascio, now at the Illinois Psychiatric Institute, Leroy Wolins at Iowa State.

The research directed its attention to methodology and policy issues. It focused on reanalysis of existing data - surveys, program evaluations, and so forth - as a vehicle for learning how to improve methods of obtaining and analyzing data, designing program evaluations, constructing policy to enhance the quality of statistical data and its utility.

Most products are used by several audiences:

- Researchers in education, sociology, psychology, political science and administration
- Public administrators with responsibility for evaluation
- Congressional and state legislative with responsibility for program evaluation law
- State and large school districts that maintain Offices of Research and/or Evaluation.

The products are relevant to practitioners such as teachers, insofar as the quality of evidence bearing on education programs can be improved by using the methods and policies we have developed.

III. Research Products and Their Uses
National Institute of Education Project on Secondary Analysis

Robert F. Boruch

This summarizes the uses of research generated by Northwestern University's Division of Methods and Evaluation Research with support of the National Institute of Education.

The main purpose is to make plain the nature of the research and how the products eventually are used by public and private agencies. All of the claims for use here can be corroborated.

1. Program Evaluation: Methods

Northwestern's work in this area is dedicated primarily to understanding how to better estimate the effects of changes in social and educational programs and of the reduction of programs. A parallel effort is dedicated to understanding how to better assess program and project implementation. It has covered the following broad problem areas:

- Scientific problems: Designing better, less equivocal ways of estimating program effects;
- Management problems: Determining how well evaluations are done and how to get evaluations done better;
- Policy problems: Developing evaluation policy at the local, state, and federal levels;
- Ethical and legal problems: Resolving problems engendered by evaluations, developing model statutes, and legal research.

The basic idea is to get better information used more often in the policy process and in research. The evaluative information here includes administrative records, surveys, and large-scale experiments. The data may be numerical or non-numerical.

The products of this research have been used internationally and domestically. Reports on the use of randomized field tests for planning and evaluating development projects in developing countries have been prepared for the U.S. Agency for International Development and presented to mission directors in Africa. This work covers methodological and management issues in evaluation; there is an extended bibliography on methodology.

Guidebooks based on the research have been published by UNESCO's Division of Social Sciences and Applications. Their special focus is on case studies of remarkably high quality evaluations (excerpts have been published in Solomon, E. S. (Ed.), Evaluating social action projects. Paris: UNESCO, 1980). Northwestern staff were the only U.S. representatives selected for recent UNESCO conferences in Malaysia on evaluation in the Far East.

In the United States, investigatory agencies such as the U.S. General Accounting Office have used the work. At GAO, for example, our guidelines on assessing the quality of federal impact evaluations, developed with foundation support, have been adapted to GAO purposes and have been published in their "Guidelines on Impact Evaluation."

Broad work on evaluation policy has been incorporated into two 1980 Reports to the Congress and the U.S. Department of Education on evaluation at the federal, state, and local levels of government, one report having been produced by the National Academy of Sciences and the other by Northwestern University. The same work has been cited in reports to the Office of Technology Assessment by Rand Corporation in special studies on educational evaluation in the United States. The Northwestern report and its recommendations were presented to the Institute for Program Evaluation at GAO in order to determine how GAO can exploit the research.

One of the policy papers generated in this series, "Implications of Educational Program Evaluations for Evaluation Policy," published in Review of Educational Research has won the research review award for 1980 from the American Educational Research Association." It was used as a basis for designing the Holtzman Project, a Congressionally mandated study.

2. Secondary Anlaysis, Criticism and Data Access Policy and Practices

One focus of this work has been on assuring that records are available for use in federal program evaluation and related applied social research. One function of access is permitting reanalysis of earlier studies.

The most recent monograph on the topic produced primarily with NIE support is: Boruch, R. F., Wortman, P. M., and Cordray, D. S. (Eds.). Secondary analysis and program evaluation. San Francisco: Jossey-Bass, 1981. This volume expands earlier work including special journal issues such as: Boruch, R. F. (Ed.). Secondary anlaysis. New Directions in Program Evaluations. Volume 4, 1979, produced with NSF and NIE support. A specialized monograph covering statistical and inferential errors has also been produced by Leroy Wolins of Iowa State University, supported by our foundation grants: A Critical Commentary on Research in the Behavioral and Social Sciences (Ames: Iowa State University Press, 1982).

The work has been international in scope and utility. For instance, we have been invited to present papers on statutory obstacles and facilitating mechanisms for reanalysis at Bielfeld and the University of Cologne, Germany. Aspects of the work were in the 1980 Israel-U.S. Conferences on evaluation and Israel's Ministry of Education and Culture may adopt some of our recommendations.

The research has been used by a variety of federal agencies. The U.S. General Accounting Office for instance has capitalized on the reports and recognized the need for competing analyses of federal evaluations in its "Guidelines on Impact Evaluation." Recommendations to the Congress in the Holtzman Report by the National Academy of Sciences Committee on Program Evaluation suggested that routine critique and secondary analysis for major educational evaluations be made a formal policy by the Department of Education and incorporated into law by Congress. The issue of critique of data and evaluations is especially crucial if block grants give states more control of evaluation. Some states are very able, other are incompetent.

Our special review of the National Center for Educational Statistics' Condition of Education series, a review commissioned by the National Academy of Education, has been instrumental in changing NCES policy on the series. As a result of formal presentations to the NCES National Advisory Board and to the Director of NCES, we have been told that modifications will be made to the series graphs, glossaries, and content.

3. Sharing Educational Research Information

Part of our work has been dedicated to assuring that information which should be accessible for research is indeed accessible. NIE and NSF have supported most of the work on the topic. The work includes development of procedures and guidelines for making information accessible, analyses of law bearing on access, and identifying obstacles to access through empirical research.

Much of the work is in the interest of improving scientific productivity and quality and to making sure that researchers can analyze data for scientific or policy purposes. Most of the work has focused on access to numerical (statistical) information on individuals. More recent work focuses on visual, narrative, and other non-numerical data.

The work has been of some interest outside the U.S. So for example, we contributed to the so-called Bellagio Principles on access to government data, based on the research. The Principles were developed at a meeting of representatives of six industrialized countries to reach a consensus on ground rules for accessing census, social security, economic and other data without breaching privacy law. The Principles have been published in at least a dozen journals and books here in the U.S. and abroad. They serve as a basis for access principles of the European Science Foundation.

Our more recent work is being used to feed into work by the National Academy of Sciences Panel on Data Sharing (Committee on National Statistics) and the Bureau of Justice Statistics (U.S. Department of Justice) Project on Access to state and local criminal records.

4. Research Policy and Management

The studies undertaken with NSF and NIE support bear on the development of research policy in the United States and abroad. It is difficult to determine exactly how the work has been used by decision-makers in this arena, but it is easy to lay out evidence that advice based on the work has been invited and presented in a variety of ways to individuals responsible for research policy or management.

Policy decisions have been informed partly through pro bono work, for instance, that involves translating our research findings into a form that addresses specific questions. We have been asked, for example, to develop memoranda for the Social Science Research Council on a variety of topics and the information has been incorporated into white papers and the like developed by SSRC to guide its activity over the last five years. The questions addressed bear on design of research during periods of scarce resources in 1981, alteration and research on social indicator systems for the United States in 1980, and social statistical research in 1979.

Similarly, formal contributions have been invited and they have been made to inform deliberations of the National Academy of Sciences Subcommittee on Data Sharing (Committee on National Statistics), the NAS Panel on Research on Quality of Work Life Experiments, the NAS Committee on Privacy and Confidentiality as Factors in Survey Response, and the Committee on Post Performance Evaluation of Basic Research. All of these contributions depended in a non-trivial way on the work supported by NIE and NSF at Northwestern University.

Similarly, we have been asked to contribute formally to the Aspen Institute's deliberations on policy for criminal justice research in Berlin and to the development of U.S. General Accounting Office policy on privacy and confidentiality of research data and policy on quality of evaluative research and have done this as well. More commonly, the contributions are a product of appointment of the principal investigator or staff to policy advisory boards. This includes participation in boards for the National Center for Educational Statistics' National Longitudinal Surveys, the GAO's development of an Institute for Program Analysis, the Social Security Administration's Cost Reduction Experiments, the Department of Health and Human Services' Seattle and Denver Income Maintenance Experiments, and others. Finally, it includes formal contribution as board members to development of the Joint Guidelines on Educational Evaluation by the American Educational Research Association and ten other cooperating institutions, to position papers elicited by the American Psychological Association on privacy and confidentiality.

5. Training

Any university-based research enterprise worthy of the label is also dedicated to advanced training. Northwestern's activity is no exception and both NIE and NSF have supported the research in which graduate students gain experience or capitalize on the experience they've accumulated elsewhere. Apart from this, specialized post doctoral programs supported by the National Institute of Mental Health and the National Center for Health Services Research have been dedicated to updating the skills of managers and researchers who have had substantial field experience.

So, for instance, former students and Fellows have gone on to reasonably important positions in academe, government and commercial sector. The government agencies to which post doctoral or predoctoral students have contributed as regular staff or managers include the Inspector General of the Department of Health and Human Services, the General Accounting Office, the National Institute of Education, the Federal Judicial Center (Research Arm of the Supreme Court), Oak Ridge National Laboratory and others. Former students hold or have held academic positions at Wheaton College, University of Colorado, George Washington University, C. W. Post University, Tennessee, City University of New York, University of West Virginia, Cornell University, and at institutions in Colombia, Norway, and Germany.

In the commercial arena, Northwestern has served as a training ground for individuals at Educational Testing Service, Abt Associates, NTS Corporation, and others.

6. Solutions to Ethical and Legal Problems in Research

Our foundation supported work on this topic is solution oriented. Attention has been dedicated partly to the problems of minimizing privacy degradation in research, assuring confidentiality of research records on identifiable respondents, and balancing needs for access to information in archives against privacy regulations, custom, or law governing archives.

Some of the recent work has been consolidated in a monograph: Boruch, R. F. and Cecil, J. S. Assuring confidentiality of social research data (Philadelphia: University of Pennsylvania Press, 1979). The monograph and our related published papers attend to statutory, statistical and procedural approaches to resolving problems. Related monographs now in press include Boruch, Ross, and Cecil's Solutions to Ethical and Legal Problems in Research (Academic, 1982) and Boruch and others' Data Sharing in Research. Each addresses other ethical and legal problems in research and explores solutions to them.

Work on the topic has been adopted by the Social Research Council's Committee on Program Evaluation in advising the U.S. General Accounting Office and is cited in both SSRC and GAO reports. It has been provided in testimony to the Privacy Protection Study Commission, was used and is cited by that group's report. Material has been adapted by the Panel on

Privacy and Confidentiality as Factors in Survey Response of the National Academy of Sciences Committee on National Statistics, and the American Psychological Association's Committee on Privacy and Confidentiality and cited in their reports.

Some of the published material has been used repeatedly over the last two years in workshops for members of federally required Institutional Review Boards that rule on ethics of research in universities. These workshops have been sponsored by the Office of Protection from Research Risks of the National Institutes of Health, the Association for Public Responsibility in Medicine and Research, and others.

The work has been recognized and used abroad as well. So for instance results of the research have been presented at Bielfeld and Cologne at the invitation of the organizors of these Germany sponsored conferences, at Stockholm for conferences sponsored by the Swedish Council for Social Science Research, and at Warsaw for the International Statistical Institute meetings. The work has been published in English and German in international professional journals, proceedings of these conferences and specialized monographs developed in Great Britain and Germany dedicated to information management and policy.

IV. Third Year Report (1981-82)

1. Secondary Analysis, Meta-analysis and Criticism

Irish Study

Reanalysis of this national experiment confirmed most of the original findings that effects of introducing standardized testing on average performance are negligible. It differs in showing remarkable increases in variance of performance and providing evidence that feedback does have a notable effect on children. Reports have been published by Locascio in Educational Measurement, Educational and Psychological Measurement and elsewhere.

Cali Study

Reanalysis confirmed original Science article findings about program effects on pre-schoolers. Analysis of three year follow-up data confirms effect but differences are diminished .6 SD. Effects are confirmed to nonschool outcomes such as reasoning, depend on duration of treatment, and do not offset entirely the advantage that high SES children have.

Title I

Meta-anlayses of over 600 norm reference and 200 regression discontinuity evaluations (Florida) suggest strongly that norm reference estimates of the effect of Title I are biased upwards, i.e., make the program look better than it is, and regression discontinuity estimates are biased downward. The norm reference estimates are plausibly biased on account of time of testing, regression artifacts, and attrition. The regression-discontinuity problems include measurement and data collection. Results are published in Evaluation Quarterly.

Condition of Education

Our published critique (National Academy of Education) of the annual series exploits standards for numbers, prose and pictures. It concludes that CE does what it is supposed to do in most respects but can be improved with respect to pictures, i.e., better graphs, prose, i.e., less naive occasional editorializing and more links to policy, and numbers, i.e., quality control information. Action has been taken on the basis of the report.

Coleman, Hoffer, Kilgore

We deliberately avoided reanalysis for conflict of interest reasons focused on the process of releasing controversial policy relevant data. The critique suggests that every such project should have a supported parallel or secondary analysis at best, sequential secondary analysis or critique at least. We have generated four policy papers on the topic however, and they are under review by various journals.

Others

See reviews by Sebring and Boruch of Wirtz and LaPointe's recommendations on the National Assessment of Educational Progress, Locascio's paper on Lord's Paradox in EPM, Wolins' book Mistakes in Research, and other products.

2. Improving Methods of Research

Randomized Experiments

The current work builds on earlier findings on criteria for judging feasibility and appropriateness to understand how experiments can be authorized under the law, how they can be coupled to nonrandomized trials, how their sensitivity and power can be increased. The corruption of our recommendations about authorization are of special interest.

Regression-Discontinuity

The Trochim reports, published in Evaluation Review and elsewhere, on Florida Title I data suggests that discrepancies between regression-discontinuity estimates and norm referenced estimates of effect are substantial and cannot yet be reconciled. New methods by Spiegleman and Trochim may reduce some problems in RD estimation.

Cross-lagged Panel Analysis

Rogosa's criticism of the method is correct. But Locascio's analysis published in EPM suggests the criticism is overdrawn: the method may be robust against moderate violation of assumptions but its robustness is still not well-understood.

Implementation

The first of two efforts in this arena is to establish better mathematical links among concepts of measuring implementation of programs, response variable irrelevance, and statistical power in evaluation design. Results are laid out in three papers, one by Boruch and Gomez has been reprinted in three different volumes, including Evaluation Studies Review Annual. The second effort is a review of over 60 randomized tests in education, showing that only a third produce any real information on implementation level, and in this about half have implementation levels in the 500-70% range. See structural models also.

Structural Models

The main completed product is a summary of the importance and triviality of structural models, written with special attention to gender-related social change and sex equity issues. The paper was published in a volume edited by Wirtenberg and Richardson. The main conclusions are that the safest applications are still infrequent and ought to be exploited better in measurement theory, randomized tests and experimental psychology. The incomplete products bear on use of models in randomized tests.

Mistakes in Research

Leroy Wolins undertook reviews of published articles in research and evaluation to identify chronic errors in statistical analyses. The latter includes problems in power, regression to the mean, scale dependent mistakes, unit of analysis errors and others. The errors appear in textbooks, including advanced level ones, and articles on IQ and environment, learning, discrimination and other areas. The book Mistakes... has been published by Iowa State University Press.

Ethical and Legal Problems

We have focused solely on disseminating earlier work in a variety of workshops, papers, and books. The main stress is on solutions, mathematical, managerial, and statutory, to problems engendered by social research in any context, including education. The published reviews of the Boruch-Cecil volume maintain that no Institutional Review Board should be without one. The two books in this series were published by University of Pennsylvania Press and Academic Press.

Evidential Problems Generally

The attitudes of congressional staff and others toward evidence in the social sector is alarmingly ignorant. Our recent efforts have focused on parallels in evidential problems among science, engineering, medicine, and educational and other social research. The problems we've examined bear on likelihood of success and failure, validity of evidence, vernacular criticism fair and otherwise, and other issues. The relevant paper was published in Studies in Educational Evaluation.

3. Research Policy Standards

Award Paper

The U.S. lacks coherent evaluation policy. The Boruch-Wortman paper, which won the AERA's research review award and published in Review of Research in Education, synthesizes a good deal of the evaluation experience of the 1970's to identify broad essential elements of evaluation policy and present evidence on what we've learned about them.

Report to Congress

We capitalized on the award paper and field research to develop recommendations to the Department of Education and Congress on evaluation of programs at the federal, state and local levels of government. That report supported by OE also identified gaps in the earlier work and have led to NIE-supported publications on questions addressed in evaluation, how well evaluations are done and who uses the results of evaluation. Relevant papers appear in Educational Researcher, American Statistician, Journal of Management and Policy, and Evaluation Review.

Title I Oversight

This is an exploratory study dedicated to understanding more about intensity of oversight of Title I. The main findings are that intensity of this oversight is considerably greater than intensity in other areas and influenced heavily by executive-Congress relations, committee interest in oversight and relations between congressional committees.

Randomized Experiments

The role that experiments plays or can play is determined partly through law and evaluation policy. The recent papers stress statutory authorization for experiments and the way appeals for authorization are corrupted, the empirical and conceptual tests necessary for randomized field studies, the failure rate of experiments and other issues. Relevant papers appear in American Statistician and Public Administration.

Others

We are currently involved in more intensive studies of differences between categorical and block grant approaches to oversight, especially in the use of information, policy that bears on use of the National Assessment of Educational Progress, policy on cooperative linkages among studies in the federal government, parallels between U.S. and Israeli policy on evaluation, and others.

4. Utilication of Research

General

A variety of our work bears on using the product of educational research and development. This includes developing evidence on use of evaluations at the federal level, especially the NIE Compensatory Education Study, on the use of methods such as regression-discontinuity and randomization, on the use of information in oversight hearings and so on. The results are published in Evaluation Review, Journal of Management and Policy, and Studies in Educational Evaluation.

NAEP

The most recent and policy-related effort on these lines is a series of ca e studies on local and state use of the National Assessment of Educational Progress. It is dramatic in presenting evidence against bald and hostile claims that NAEP is not used; it is imperfect in that resources did not permit statistical characterization of use based on probability samples. The used products include methods, rather than just data. Recommendations were discussed at NAEP policy board meetings. The work is published in Educational Measurement: Issues and Practice, an AERA journal.

V. Second Year Report and Abstracts: 1980-81

1. Introduction

This is the second annual report on research supported by NIE-G-79-0128. The grant runs three years, subject to renewal at approximately $60,000 per year at Northwestern University, (Boruch as P.I.) and $70,000 per year at University of Michigan (Wortman as Co-PI).

The first annual report covers predecessors to this grant and the first year's products. This report provides updated information, including revisions to the first year's work. Brand new papers not mentioned in the first annual report are indicated by an asterisk.

A separate progress report has been prepared by Wortman.

2. Research on Program Implementation

Our main interest lies in how to measure program implementation, how to characterize implementation numerically, and how to exploit resulting data in analysis.

Two papers illustrate the work.

→ Thomas, S. and Boruch, R.F. Numerical indicators of program implementation in field tests: Education. Working Paper A-91. Division of Methodology and Evaluation Research; Psychology Department, Northwestern University, Evanston, Illinois, March, 1980. (1)

This paper reports an effort to assay implementation of programs tested in over 60 field experiments. The target population was the listing of education experiments given in Boruch, McSweeny, and Soderstrom's bibliography on randomized trials. Of these, only 13 of the reports on the experiments carried sufficient information to assay level of program implementation. A bit less than half the programs appear to have been implemented to 50-70% of their ideal value. The remainder exceed 70% implementation levels. Indices of implementation vary considerably from one study to the next and include p.

→ Thomas, S. Numerical characterization of program implementation. MS Thesis, Psychology Department, Northwestern University, Evanston, Illinois, May, 1981. (2)*

This thesis is dedicated to consolidating available numerical information on level of educational program evaluations. No other synthesis of such data has yet been undertaken and the report should be of some interest to program managers and developers, and to evaluators. The generic differences between media based implementation (TV and Radio) and class based programs, and their implications for measuring implementation will be laid out. Some theoretical work will be done to organize the hierarchy of measurement efforts, ranging from (say) macroassessment of structure down to micorassessment of children's reception of information. We expect that there are generic changes in implementation over time, and we hope to identify these as well. The work extends earlier work by Boruch and Wortman and by Boruch and Gomez, supported by NIE. It was supervised by Boruch and Cordray.

3. Research on Special Evaluation Designs

So-called regression discontinuity designs are a promising approach to estimating the effects of programs when individuals' assignment to the program is based on measurable need. Title I program regulations suggest this as one of the designs to be used at the local level for program evaluation. No work appears to have been done, aside from an early 1974 effort by Boruch and DeGracie in Mesa, to understand problems in implementing the design at the local level.

Since September 1979, William Trochim has been obtaining data from LEAs that have adopted the design and has been interviewing evaluators who have tried to use it. The products of that work have been consolidated in a dissertation:

→ Trochim, W. K. Regression-discontinuity designs in Title I program evaluation. Ph.D. Dissertation, Department of Psychology, Northwestern University, Evanston, Illinois, August 1980. (3)

Chapters of the dissertation are being issued independent reports. The following is illustrative. One chapter deals with reanalysis of data generated at the local level and it is described briefly in the section entitled Secondary Analysis of Specific Data Sets in this report.

→ Trochim, W. M. K. Implementation of regression discontinuity designs in Title I program evaluation. Report No. T-1. Division of Methodology and Evaluation Research, Psychology Department, Northwestern University, Evanston, Illinois, July 1980. (Chapter II, of Ph.D. Dissertation of the same title and origin). (4)

This report explores problems in implementing the regression discontinuity (RD) design for evaluating Title I programs. Data stem from 50 interviews with SEA and LEA evaluators and Technical Assistance Center personnel, from reanalysis of RD data generated by school districts in Rhode Island and Florida, and from other sources.

According to recent NCES surveys, the large majority of school districts (86%) use the weakest approach to estimating Title I programs. Less than 2% use regression-discontiniuty approaches; less than 20% use other control group approaches. The largest proportion of R-D design users are in Florida, New Jersey, and Virginia.

Our interview data on RD suggest the following as reasons why the design is not used: (1) it is likely to yield negative estimates of effect (norm referenced approaches give high positive estimates), (2) the weaker norm referenced approach is more "familiar" to school districts, (3) the RD requires a cutoff to identify who is neediest and some administrators want more discretion, (4) the RD design based on need with lead to loss of teachers, (5) the SEA discourages the model's use, (6) the early results on RD suggested that it would give negative results, (7) RD is technically more difficult than the weaker approach, (8) RD requires testing of comparison group-students, (9) RD requires at least 30 students in the treatment group, (1) RD requires a pre-post correlation of .4 according to the OE policy manual.

The meta-analyses done on over 25% conventional RD applications by the Florida Department of Education confirm that the estimates of effect are on average negative. In Florida, NCE gains average 46 NCE units. The author believes both types of estimators are biased but in different directions. The reasons for the negative estimates may be: misspecification of the underlying model (nonlinear instead of linear, the latter being assumed in Title I documentation). Note that the numbers of students problem here is legitimate. If 185,000 are served in Title I and 15,000 school districts have the program, this means only 12 or so kids per district are served. If this is the case, many analyses will not be sensitive enough through the "30 in each group" rule is probably silly. Misassignment challenges are another source of problem in implementing design, but no hard data on misassignment or challenges is available. Ceiling and floor efforts may produce nonlinearity and there is some evidence from Florida that this is indeed the case.

Other chronic problems include the difficulty of verifying what the nature of instruction was, amount of instruction. The exclusion of students from the program presents major problems for analysis. This includes attrition, about 30% in Florida, and nobody knows what to do about it. "Common" deliberate exclusions include students with incomplete data.

→ Trochim, W. K. and Spiegelman, C. The relative assignment variable approach to solution bias in pretest posttest group design. Proceedings of the Survey Research Methods Section, American Statistical Association, 1980, Washington, D.C.: ASA, 1981, pp. 376-381. *

The paper describes the mathematics underlying regression discontinuity designs and a procedure for estimating parameters, notably size of program effects, when the design is used.

→ Boruch, R. F. Social experiments in education. Proceedings of Survey Research Methods Section, American Statistical Association, 1980, Washington, D.C.: ASA, 1981, pp. 388-393. *

Revised and accepted for publication as: Boruch, R. F. Experimental tests in education: Recommendation is from the Holtzman Report. American Statistician, 1982 (February), in press. *

Extended, to appear as Boruch, R. F. Randomized experiments at national, state, and local levels of government: Recommendations of a Report to Congress. In C. Aslanian (Ed.) Proceedings of the Evaluation Research Society, 1982, in press. *

The paper covers recommendations to the Congress and the U. S. Department of Education bearing on experimental tests of programs to improve education or decrease costs. The discussion includes conditions under which experiments are feasible and appropriate, constraints, and areas where research on the topic is needed.

The American Statistician version includes commentary and criticism from statisticians.

→ Locascio, J. J. The cross-lagged correlation technique: Reconsideration in terms of exploratory utility, assumption specification, and robustness. Submitted to and under review by Educational and Psychological Measurement (1981). *

The cross-lagged correlation technique (CLC) for assessing causality from passive observational data has been the subject of much recent criticism. CLC compares cross correlations between variables across time points of measurement and attributes differences in correlations to causal effects. Some of the major areas of criticism have to do with: (a) lack of a no-cause baseline, (b) spurious effects of mediating variables, and (c) obscuring effects of heterogeneous stabilities. The present author attempts to show that while these criticisms are valid, CLC should not be rejected outright as a tool in social science research. CLC has some utility as an exploratory technique. Furthermore, if certain assumptions are specified and these assumptions are correct, valid inferences can be made from CLC. Lastly, some assumptions for CLC are not as stringent as implied by critics and CLC is robust to minor violations of assumptions.

→ Locascio, J. J. and Cordray, D. S. A reanalysis of Lord's paradox. * Unpublished manuscript. To be submitted to Educational and Psychological Measurements.

"Lord's Paradox" is a term applied to a hypothetical stiuation illustrated by Frederick M. Lord in which Gain Score Analysis (GSA) and the Analysis of Covariance (ANCOVA) yield conflicting results in testing differential treatment effects for natural groups. The authors further analyze this paradox and demonstrate that GSA and ANCOVA will always yield conflicting results unless the slopes of within group regression lines of posttest on pretest equal one. It is also pointed out that Lord's specific hypothetical example, trait instability accounts for regression line slopes not equalling one and invalidates the ANCOVA. Recommendations for analysis of pre/post data for preexisting natural groups are provided.

→ Locascio, J. J. Problems in the exploitation of extra-agency based * management information systems for use in intra-agency program monitoring and evaluation. Submitted to and under review by Evaluation Review (1981).

A social service agency may to a large extent be able to monitor and evaluate its own programs via an extra-agency based management information system (MIS) into which it reports, thus avoiding the cost and additional labor required to run its own information system. However, some adjunctive information systems internal to the MIS reporting agency may be needed to provide raw data not provided by the MIS, collation schemes which are more useful to the reporting agency, and/or speedier processing of some information than may be provided by the MIS. Such adjunctive systems may be economically tied into the MIS rather than being completely independent of it. Adjunctive systems should be flexible enough to be able to supply information previously supplied through the MIS but lost through unforseen modifications in the MIS by the source agency. Accuracy of reported data may be an especially troublesome problem when using large extra-agency based MISs because of complicated reporting procedures. Reporting error should be constantly monitored and action taken to keep error rates minimal.

4. Secondary Analysis of Specific Data Sets

The Project's emphasis based on the premises that secondary analysis can be helptul in establishing credibility of original evaluations, extending theory and advancing methods, and in testing new methods. The choice of particular data sets for this depends on (a) importance of the data set, (b) availability of interested staff to work on reanalysis, and (c) accessibility of the data and quality of its documentation.

Reanalyses that are currently underway or complete include the following items. Note that the purpose of the exercise differs from one data set to the next.

→ McKay, A. C. The long term effect of a multiform program of preschool intervention on the cognitive abilities of chronically deprived children. Ph.D. Dissertation. Psychology Department, Northwestern University, Evanston, Illinois 1981 (in progress).

This is a reanalysis and extension of analyses of research data on cognitive abilities generated by randomized field experiments in Cali, Colombia. Professor Robert Boruch is serving as chairman of a Dissertation committee which includes Professors David Cordray of the Psychology Department and David Wiley and Bruce Spencer of the School of Education. The dissertation will be completed no later than December, 1980.

The topic of the dissertation is "The long-term effect of a multiform program of preschool intervention upon the cognitive abilities of chronically deprived children." It is based on data collected during the period from 1974 to 1978, from a randomized experiment designed to assess the long-term effects of the preschool intervention carried out between 1971 and 1974 in Cali, Colombia. Our article, published in Science in 1978, "Cognitive Ability in Chronically Deprived Children," presented the major findings of the preschool intervention period and results of one I.Q. measure, The Stanford-Binet, applied one year after treatment ended. Since that time there has been no other published report concerning the follow-up research undertaken.

The scope of the dissertation will be limited to one aspect of the many for which data are available: cognitive abilities. These results are of the highest priority when seen in the context of the findings from the preschool years and in light of present research issues in early childhood education in the United States.

Historically the Cali project has been of importance because of many unique features including: (1) it was a large scale randomized field experiment with combinations of treatments systematically varied across groups; treatments that until that time had never been applied simultaneously; (2) the subject population was living at levels of physical and economic deprivation beyond any reported in United States early childhood development studies. Although research on the effects of malnutrition on cognitive abilities had included severely deprived populations, never had these studies utilized education or behavioral intervention as a treatment variable; and (3) a careful control for measurement history by applying all measures of cognitive abilities on all experimental groups of subjects at all testing periods regardless of the time of entrance of subjects into treatment. Thus, over a seven year period, subject data is comparable and permits groups awaiting treatment to serve as controls.

These factors continue to be important as the issues of long-term impact are addressed. In the United States, initial evaluations of long-term impact were relatively negative. Data seemed to indicate that once children entered the mainstream of education the sometimes dramatic gains in intellectural functioning disappeared and after one or two years there were no significant differences between treated children and controls. Many methodological problems as well as substantive issues have confused the interpretations of findings, however, and thre is an emerging body of analyses that suggests that these initial findings do not really reflect the effects of intervention.

A major research effort, the Developmental Continuity Consortium Study, is underway to reanalyze data from longitudinal studies as well as collect and analyze new data on children who were subjects of intervention studies in the late 1960s and early 1970s. The decision of these researchers was to apply one major intelligence test, The Wechsler Intelligence Scale for Children, plus measures of school performance. From its inception the Cali project employed as part of its test battery some Wechsler measures. This means that, subject to limitations, the Cali results can be compared with forthcoming U.S. data and should have significant influence upon interpretations in this problematic area.

The dissertation will encompass a review of all relevant and important studies reported in the past 15 years, a description of the study, its treatments and measures, and data analyses sufficient to provide answers to four basic questions widely and presently asked about the long-term effects of preschool experience:

1. Does the effect of preschool intervention maintain itself over the first three years of primary school?

2. Is the persistence of effects a function of the amount of treatment during the preschool years?

3. Are separate, specific, cognitive abilities affected differently by treatment?

4. Is the amount of treatment related to school grade placement?

The major independent variable to be studied is duration of preschool treatment: 4,3,2, or 1 treatment periods prior to school entrance. An untreated comparison group (0 treatment periods) of children from the same slum neighborhood, but eliminated from intervention because their health status was normal will be included in the analyses, as they were in the Science article. The high social status comparison group, included in the Science article, will be dropped from this set of analyses.

The dependent variables will include (1) cognitive measures; specifically, the Stanford-Binet intelligence test and individual subscales of the Wechsler Intelligence Scale for Children-Revised, and (2) grade placement at the end of three years.

Data will be analyzed for effect of treatment at two measurement periods: (1) after the conclusion of the preschool intervention period, and (2) after the conclusion of three follow-up years; for children at correct grade placement this would mean after the third grade of primary school. Finally, although low, subject attrition will be analyzed.

→ Trochim, W.M.K. Statistical analysis of the regression-discontinuity design. Report T-2. Division of Methodology and Evaluation Research, Psychology Department, Northwestern University, Evanston, Illinois, July 1980 (Chapter III of Ph.D. Dissertation by the same author).

This report examines a strategy of "overfitting" data in sharp cutting point regression-discontinuity designs. That is, one posits that most curves underlying RD data are quartic and perhaps cubic rather than linear, and fits the higher order polynomial. The major reason for fitting beyond the linear even where linear is a least squares best fit is that (a) overfitting leads to no bias in estimates and to negligible loss in power, and (b) fitting only linear functions leads to biased estimators, in the author's judgment. Computer simulations are used to illustrate both points. And data from three school districts (twenty-four regression-discontinuities) are used as a vehicle for understanding the worth of the strategy.

"Overfitting" is something of a misnomer here. The main idea is to fit higher order terms routinely because the author believes that the underlying curves are indeed higher order polynomials and the overfitting is safer than merely fitting linear functions.

→ Ross-Reynolds, J. Examination of item bias on the WISC-R for four socio-cultural groups. Thesis. Department of Psychology, Iowa State University, Ames, Iowa, 1980.

This is one of a very small number of works which attempt to assay test item bias in the WISC-R tests. The incentives for the investigation stem from often vague and unsubstantiated allegations that conventional achievement tests are culture biased, so making judgements about individual or group ability inaccurate and making judgements about the effectiveness of ameliorative programs risky. The research examines test item difficulty across four socio-cultural groups: Anglo, Black, Chicano, and Native American Papago (Indian). Data for the study come from research originally undertaken by Reschly and Jipson in Pima County Arizona. Methods used to reanalyze these data include computation of p-values (item pass rates), transformed item difficulties, tests and analysis of outliers, and point biserial correlation between item performance and raw subtest score. The evidence does not support the claim that item difficulties vary notably with social-cultural group, except perhaps for the Papagos. That is, items are appreciably more difficult for Blacks or Chicanos than they are for Anglos. This work examines only one aspect of test bias. Biases engendered by incompetent use of tests, for instance, is ignored. It does review recent court decisions on test bias, and reviews recent related statistical research on the topic.

→ Locascio, J. and Boruch, R. F.

We are planning to obtain the Irish Study for reanalysis. We have some data from early NIE career education programs and are trying to figure out whether it should be and can be reanalyzed.

→ Wolins, L. Our teaching mistakes lead to mistakes in research. In Proceedings of the American Statistical Association: Social Statistics Section, 1980. Washington, D.C.: ASA, 1981, pp. 386-387. *

In the social-behavioral science literature are many instances of incorrect or inefficient design and analysis of even simple two-treatment or two-group experiments and investigations. Some of the responsibility for these mistakes can be attributed to textbook writers and how we teach statistics. We tend to advocate equal sample sizes in randomized group experiments in the face of unequal variances or unequal cost in administrating treatments. We offer Wilcoxson's Rank-Sum Test only for randomized groups even though it is appropriate for paired observations and more powerful than the sign-test. We fail to inform our students that Wilcoxson's Signed-Rank Test is scale dependent and thereby usually inappropriate for basically the same reason that the t-test is inappropriate and, finally, a negative rank-order correlation for paired observations is as devastating for making proper inferences based on non-parametric analysis as is the comparable kind of non-additivity in parametric analysis. These and other errors are discussed. This work is based on review of recent articles appearing in a dozen journals such as Journal of Educational Psychology, Psychological Bulletin, and others. It is a by-product of an effort to document and discuss such errors in monograph form.

→ Wolins, L. A critical commentary on research in the social and behavioral sciences. Ames: Iowa State University Press, 1982, in press. *

The volume has the following chapter contents.

1. Power, Sufficiency, and Admissibility Problems
2. Regression toward the Mean Mistakes
3. Scale Dependent Mistakes
4. Variable and Observational Unit Mistakes
5. Data Exploitation and Exploration Mistakes and Techniques
6. Instrument Development: Problems, Admonitions and Some Partial Solutions

A paper that focuses on some of these matters was developed for the Boruch, Wortman, Cordray volume described later.

→ Trochim, W. M. K. and Spiegelman, C. The relative assignment variable approach to selection bias in pretest-posttest group designs. In: Proceedings of the American Statistical Association: Social Statistics Section. Washington, D.C.: ASA, 1981, pp. 376-381.

Pretest-posttest group designs are a class of research methods widely used in evaluation and applied social research. These designs can be classified by the degree to which the process of selection into groups is specified. With the true experiment and regression-discontinuity designs the selection process is known and unbiased estimates of treatment effect can be obtained using OLS. With the non-equivalent group design the selection process is seldom, if ever, known. Recent work suggests that the selction bias which occurs in the cases outlined above can be accounted for by the estimation of a new variable which reflects the "relative assignment" or the proportion of participants in the treated condition for every pretest value. Three approaches for estimating relative assignment are investigated. The illustrative data are from the State of Florida's Title I Testing Program.

5. Evaluation Policy and Metaanalytic Reviews

One of the themes of the Project is in development of better evaluation policy and related statistical policy in education. Much of this work depends heavily on review of multiple program evaluations, on critique and secondary analysis of some, and on research on statistical products and policy.

The work was begun in 1979 and has been supported by NIE, and by NSF when NIE funds were unavailable. The following papers represent current products of the effort.

→ Boruch, R. F. and Wortman, P. M. Implications of educational evaluation for evaluation policy. In: D. Berliner (Ed.) Review of Research in Education, Volume 7, Washington, D.C.: American Educational Research Association, 1979, pp. 309-360.

The aim of this paper is to review certain policy implications of evaluative research. Our special interest lies in evaluation policy, rather than in the broad category of educational policy, which may be educed from research results.

Of the several reasons for the choice of topic, not the least is a simple, perhaps morbid curiosity about what can be said about evaluation policy, taking account of recent large-scale efforts to estimate the efforts of educational programs. A second reason is rather more pragmatic. Policy does affect the quality of evaluative research (at least at times), and the United States lacks a coherent evluation policy in education as it lacks such policy in health services and other areas. We do not pretend that this paper provides coherent policy. However, we do believe the matters covered here are a legitimate part of any durable evaluation policy.

Part 1 of the paper deals with the recent history of evaluation including the instability of national interest in evidence, unfounded discouragement, and catholicity. Part 2 focuses on design of evaluations, especially its influence on the quality of information that is generated about an educational program. The implementation of new pilot programs and the independent execution of evaluation design are discussed in Part 3; policy implications of collective experience in analyzing results are discussed in Part 4. Part 5 deals with the availability of information on evaluations and the lack of it. The last major section, Part 6, covers the utilization of evaluation results.

→ Wohlstetter, P. and Boruch, R. F. Numbers, Prose, and Pictures: Review of the "Condition of Education." In J. Carroll (Ed.) Proceedings of the National Academy of Education. Washington, D.C.: NAE 1980, in press.

This paper is an extended review of the series of publications issued annually by the National Center for Educational Statistics, under the title Condition of Education. The review covers publications issued between 1975 and 1980. It focuses on the quality of the monograph graphics, the policy and theory underlying production of social indicators books such as these, and quality and character of the data. There is some attention to provincialism in choice of topics, display format, numerical rather than non-numerical indicators. The review will run to 10,000 words or more. It was invited by John Carroll, current Vice President of the National Academy of Education.

→ Boruch, R. F. Social experiments: Education. In Proceedings of the American Statistical Association Social Statistics Section. Washington, D.C.: ASA, 1981. Extended and revised and reprinted in American Statistician and Aslanian volume (see above).

In principle, randomized field experiments can provide unbiased estimates of the effects of new programs. In practice there are a number of managerial, political, and other problems which affect the feasibility and appropriateness and the quality of resultant data. The paper discusses recommendations bearing on experiments in educational program evaluation, presented to the Department of Education and the Congress in June 1980. The recommendations suggest greater use of the design, pilot tests of major new programs, program components, and program variations to facilitate their case, and explicit statutory provisions for the use of experiments.

→ Boruch, R. F., Wortman, P. M., and Cordray, D. S. (Eds.) Secondary Analysis. San Francisco: Jossey-Bass, 1981, in press.

A book manuscript, begun under our earlier NIE contract is being completed with current NIE grant support. The volume's focus is on secondary analysis in education and in other areas. It covers policy, operations, and cases. The volume will be issued as described.

A table of contents is given in the exhibit attached. Note that work on indexes for the volume and the normal drudgery of publication will continue through the second year of the project.

→ Boruch, R. F., Cordray, D. S., Pion, G., and Leviton, L. An appraisal of federally supported educational programs (Holtzman Project). Report to Congress. June 1980. Under revision for publication by Cambridge University Press, 1982.

The work for this Project was not supported by NIE. However, most of the federally supported program evaluation is relevant to the NIE project. The report addresses questions such as why evaluations are done, how well they are done (methods and standards), and how they are used. For instance, we've done case studies on utilization of every major program evaluation done over the past three years. It is sensible to build on that work and develop cost benefit appraisals of the evaluations where possible, to better document implementation issues, to assay the need for secondary analyses in these cases. We plan to use some of the NIE support, a marginal investment, to extend this work and to develop a monograph on the topic. The following publications stemming from the research were however supported by the NIE grant:

Boruch, R. F., Cordray, D. S., Pion, G., and Leviton, L. A mandated appraisal of evaluation practices: Digest of recommendations to the Congress and the Department of Education. Educational Researcher, 1981, 10, 10-13, 31. *

Boruch, R. F. and Cordray, D. S. The Northwestern report. Evaluation report. Evaluation Research Society Newsletter, 1981, 5(2), 6. *

Boruch, R. F. and Cordray, D. S. Rejoinder to Bunda and House. Evaluation Research Society Newsletter, 1981, 5(2), 8. *

Boruch, R. F. and Cordray, D. S. On criticism of the Holtzman Report. Educational Researcher, 1982, in press. *

→ Boruch, R. F. How can and should potentially controversial research results be disclosed? Presented at the NIE Conference on Methodology and Credibility of Results, Washington, D.C. July 2--22, 1981. *

The paper was prepared for a conference on conflicts among reports by Coleman, NAEP, and NIE on private and public schools. The paper is being revised for publication.

→ Boruch, R. F. On evaluation policy in the U.S. and Israel. In S. Kugelmas and A. Lewy (Eds.) Evaluation Policy and Practice. Tel Aviv, 1982, in press. *

The paper compares American and Israeli practices and policy on process and implementation studies. The paper was presetned at and developed from the joint U.S.-Israel seminars on evaluation in Jerusalem, in June, 1980.

→ Boruch, R. F., Cordray, D. S., and Pion, G. How well are evaluations carried out? Excerpts from a report to Congress and the Department of Education. In L.-E. Datta (Ed.) Proceedings of the Evaluation Research Society. Beverly Hills, Sage, 1982, in press. *

The paper extends work in the Holtzman Report on quality of evaluations at the national, state, and local levels of government. It covers empirical studies on quality, constraints on quality including manpower capabilities, federal law and regulations and grant review processes, areas where more research is needed, the utility of guidelines and related topics.

VI. Reference List

Project on Secondary Analysis

Reports of Research by Boruch
Published in Books and Scholarly Journals

BOOKS AND MONOGRAPHS

Boruch, R. F., Wortman, P. M., & Cordray, D. (Eds.) Reanalyzing program evaluations: Policies and practices for secondary analysis of social and educational programs. San Francisco: Jossey-Bass, 1981.

Wohlstetter, P., & Boruch, R. F. Numbers, prose, and pictures: A review of "The condition of Education." Proceedings of the National Academy of Education. Washington, D.C.: National Academy of Education, 1981. (Monograph).

Boruch, R. F., & Cecil, J. (Eds.). Solutions to ethical and legal problems in social research. New York: Academic Press, 1982.

Boruch, R. F., & Cordray, D. S. (Eds.). An appraisal of educational program evaluations: Federal, state, and local agencies. New York: Cambridge University Press, 1982 (in preparation.

Journal Articles
by Boruch

Boruch, R. F. and Gomez, H. Sensitivity, bias, and theory in impact evaluations. In L. Sechrest (Ed.) Evaluation Studies Review Annual, 1979, 4, pp. 126-152. (Reprint).

Boruch, R. F. and Cecil, J. S. Report from the United States: Emerging data protection and the social sciences need for access to data. In E. Mochman and P. Muller (Eds.), Data protection and social science research. Frankfurt/New York: Campus Verlag, 1979, pp. 104-128.

*Boruch, R. F. and Wortman, P. M. Some implications of educational evaluation for evaluation policy and program development. Review of Research in Education, 1979, 7, 309-361.

Boruch, R. F. Some implications of educational program evaluations for evaluation policy. American Behavioral Scientist, 1979, 23, 275-296. (Reprinted excerpts).

Hedrick, T. E., Boruch, R. F., and Ross, J. Policy and regulation for ensuring the availability of data for secondary analysis. In: L. Sechrest and Associates (Eds.), Evaluation Studies Review Annual 1979, 4, pp. 385-406. (Reprint of earlier article).

Boruch, R. F., Anderson, P., Rindskopf, D., Amidjaya, I., and Janson, D. Randomized experiments for planning and evaluating local programs: A summary on appropriateness and feasibility. Public Administration Review, 1979, 39, 36-40.

*Received the American Educational Research Association Review Award in 1980.

BEST COPY AVAILABLE

Boruch, R. F. Methods for assuring personal integrity in social research: An introduction. In M. Bulmer (Ed.), Censuses, surveys, and privacy. London: Macmillan (New York: Holmes and Meier), 1979, pp. 235-248.

Boruch, R. F. and Reis, J. The student, evaluative data, and secondary analysis. New Directions in Program Evaluation, 1979, 8, 59-72.

Boruch, R. F. and Gomez, H. Power theory in impact evaluations. In L. E. Datta and R. Perloff (Eds.), Improving evaluations. Beverly Hills, Ca.: Sage, 1979, pp. 139-170. (Expanded version of earlier article).

Boruch, R. F. and Cecil, J. S. Some solutions to privacy problems engendered by federal regulation and social custom. In M. L. Wax and J. Cassell (Ed.), Federal regulations: Ethical issues and social research. Boulder, Colorado: Westview, 1979, pp. 173-184.

Boruch, R. F. Case studies of high quality outcome evaluations. Chapter 3 of E. S. Solomon (Ed.), Evaluating social action projects: Principles, methodological aspects, and selected examples. Paris: UNESCO, 1980, pp. 105-161.

Boruch, R. F., Cordray, D. S., Pion, G. M., and Leviton, L. A mandated appraisal of evaluation practices: Digest of recommendations to the Congress and the Department of Education. Educational Researcher, 1981, 10(4), 10-13, 31.

Boruch, R. F. On evaluation policy in the United States and Israel. In A. Lewy and S. Kugelmass (Eds.) Decision oriented evaluation in education: The case of Israel. Philadelphia and Rehovot: International Science Services Press, 1981, pp. 1-32.

Boruch, R. F., Cordray, D. S., and Wortman, P. M. Secondary Analysis: Why, how, and when. In R. F. Boruch, P. M. Wortman, and D. S. Cordray and Associates. Reanalyzing program evaluations. San Francisco: Jossey-Bass, 1981, pp. 1-19.

Boruch, R. F., Cordray, D. S., and Pion, G. How well are local evaluations carried out? In L. E. Datta (Ed.), Evaluation in change: Meeting new government needs. Beverly Hills, CA.: Sage, 1981, pp. 13-40.

Boruch, R. F. Evidential problems in social program evaluation and their analogs from engineering and the physical sciences, medicine, and an assortment of other disciplines. Studies in Educational Evaluation, 1981, 7, 121-149.

Boruch, R. F. and Cordray, D. S. Reactions to criticism of the Holtzman Report. Educational Researcher, 1981, 10, 10-12.

Boruch, R. F. Randomized experiments at national, state, and local levels of government. In C. B. Aslanian (Ed.), Improving educational evaluation methods. Beverly Hills, CA.: Sage, 1981, pp. 41-66.

BEST COPY AVAILABLE

Boruch, R. F. Recommendations from the Holtzman Report: Experimental tests of education programs. American Statistician, 1982, 36, 1-7.

Boruch, R. F. and Cecil, J. S. Statistical strategies for preserving privacy in direct inquiry. In J. E. Sieber (Ed.) The ethics of social research: Surveys and experiments. New York and Heidelberg: Springer-Verlag, 1982, pp. 207-233.

Boruch, R. F. Rejoinder to Datta and Wallman on randomized experiments American Statistician, 1982, 36, 13-14.

Sebring, P. A. and Boruch, R. F. How is the National Assessment of Educational Progress used? Results of an exploratory study. Educational Measurement: Issues and Practice, 1983, 2(1), 16-20.

Boruch, R. F., Cordray, D. S., Pion, G. and Leviton, L. Recommendations to the Congress and their rationale: The Holtzman project. Evaluation Review, 1983, 7(1), 5-36.

Boruch, R. F. and Wohlstetter, P. On educational indicators. Contemporary Education Review, 1983, 2(1), 13-22.

Boruch, R. F. The administration of a Congressional mandated study at federal, state, and local levels of government. Studies in Educational Evaluation, 1983, 9(1), 89-104.

Boruch, R. F. Causal models: Their import and their triviality. In J. Wirtenberg and B. Richardson (Eds.) Methodological issues in sex roles and social change. New York: Praeger, 1983. (In press).

Leviton, L. and Boruch, R. F. Usefulness of evaluations: The NIE Compensatory Education study. Journal of Policy Analysis and Management, 1983, in press.

Leviton, L. C. and Boruch, R. F. Contribution of evaluation to education programs and policy. Evaluation Review, 1983, in press.

Publications
by Wortman

Wortman, P. M. Evaluation research: A methodological perspective. Annual Review of Psychology, 1983, 34, 223-260.

Wortman, P. M. and Yeaton, W. H. Synthesis of results in controlled trials of coronary artery bypass graft surgery. In R. J. Light (Ed.), Evaluation Studies Review Annual, 1983, 8 (Beverly Hills, CA.: Sage).

Wortman, P. M. School desegregation and black achievement: An integrative review. Sociological Methods and Research, 1983, in press.

BEST COPY AVAILABLE